Снижение размерности пространства в задачах декодирования сигналов

Исаченко Роман Владимирович

Диссертация на соискание ученой степени
кандидата физико-математических наук

05.13.17 – Теоретические основы информатики

Научный руководитель: д.ф.-м.н. В. В. Стрижов

Москва, 2021 г.

Снижение размерности пространства в задачах декодирования сигналов

Исследуется задача выбора модели при восстановлении скрытых зависимостей в исходном и в целевом пространствах.

Проблема

Целевая переменная – вектор, компоненты которого являются зависимыми. Гетерогенные пространства исходных и целевых переменных обладают существенно избыточной размерностью.

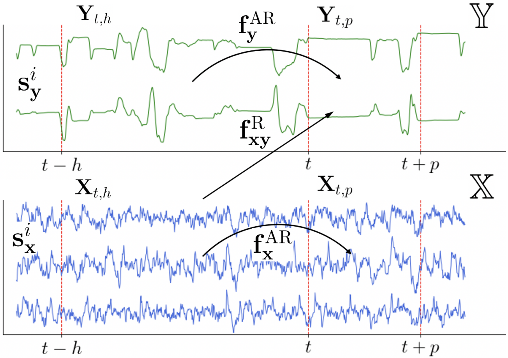
Требуется

Требуется построить модель, адекватно описывающую исходное и целевое пространства при наблюдаемой мультикорреляции в обоих пространствах.

Метод решения

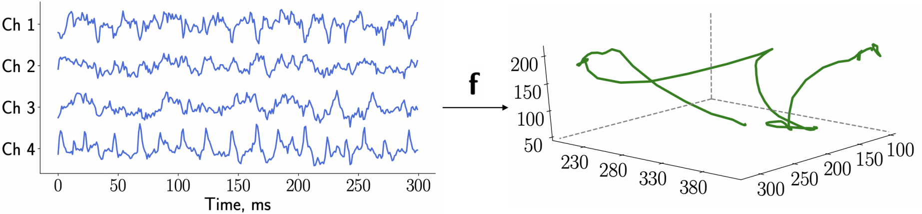
Предлагается снизить размерность путём проецирования исходных и целевых переменных в скрытое пространство. Предлагаются линейные и нелинейные методы согласования прогностических моделей в пространствах высокой размерности.

Задача декодирования сигналов

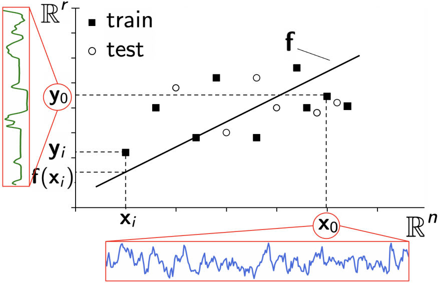


$\mathcal{S}_{\mathbf{x}} = \{\mathbf{s}_{\mathbf{x}}^{i}\}_{i=1}^{m}$ – множество временных рядов.
$\mathbf{x}_t = ([\mathbf{s}_{\mathbf{x}}^{1}]_t, \dots, [\mathbf{s}_{\mathbf{x}}^{m}]_t) \in \mathbb{R}^m$ – временное представление.
$\mathbf{X}_{t,h} = [\mathbf{x}_{t-h+1}, \dots, \mathbf{x}_t]^{\mathsf{T}} \in \mathbb{R}^{h\times m}$ – представление предыстории.
$\mathbf{X}_{t,p} = [\mathbf{x}_{t+1}, \dots, \mathbf{x}_{t+p}]^{\mathsf{T}} \in \mathbb{R}^{p\times m}$ – представление горизонта прогнозирования.
$\mathbf{f}_{\mathbf{x}}^{\mathrm{AR}} : \mathbb{R}^{h\times m} \to \mathbb{R}^{p\times m}$ – авторегрессионная модель.
$\mathbf{f}_{\mathbf{xy}}^{\mathrm{R}} : \mathbb{R}^{h\times m} \to \mathbb{R}^{p\times r}$ – регрессионная модель.
$\mathbf{f}_{\mathbf{xy}} : \mathbb{R}^{h_x\times m} \times \mathbb{R}^{h_y\times r} \to \mathbb{R}^{p\times r}$ – модель декодирования.

Восстановление зависимости в исходном и целевом пространствах



Прогностическая модель декодирования



Согласование зависимостей в скрытом пространстве

$$
\begin{tikzcd}
\mathbf{x} \in \mathbb{R}^n \arrow[rr, "\mathbf{f}"] \arrow[dr, bend right, "\mathbf{W}"'] & & \mathbf{y} \in \mathbb{R}^r \arrow[dl, bend left, "\mathbf{C}"] \\
& \mathbf{t}, \mathbf{u} \in \mathbb{R}^\ell \arrow[ul, bend right, "\mathbf{P}"'] \arrow[ur, bend left, "\mathbf{Q}"] &
\end{tikzcd}
$$

$$\mathbf{x} = \mathbf{P}\mathbf{t} + \mathbf{e}_\mathbf{x}$$

$$\mathbf{y} = \mathbf{Q}\mathbf{u} + \mathbf{e}_\mathbf{y}$$

$$\text{cov}(\mathbf{t}, \mathbf{u}) \to \max_{\mathbf{P}, \mathbf{Q}}$$

Задача декодирования сигналов

$\mathbf{Y} = \mathbf{F}(\mathbf{X}, \boldsymbol{\Theta}) + \mathbf{E}_\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\Theta}^{\mathsf{T}} + \mathbf{E}_\mathbf{y}$ – модель с параметрами $\boldsymbol{\Theta} \in \mathbb{R}^{r \times n}$.

Функция потерь модели декодирования

$$\mathcal{L}(f, \mathbf{X}, \mathbf{Y}) = \|\mathbf{Y} - \mathbf{F}(\mathbf{X}, \boldsymbol{\Theta})\|_2^2 = \Big\| \underset{m\times r}{\mathbf{Y}} - \underset{m\times n}{\mathbf{X}} \cdot \underset{r\times n}{\boldsymbol{\Theta}}^{\mathsf{T}} \Big\|_2^2 \to \min_{\boldsymbol{\Theta}}.$$

Метод проекции в скрытое пространство

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathsf{T}} + \mathbf{E}_\mathbf{x},$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^{\mathsf{T}} + \mathbf{E}_\mathbf{y}.$$

$$\begin{tikzcd}
\mathbb{X} \subset \mathbb{R}^n \arrow[r, "\mathbf{f}"] \arrow[d, shift right, "\mathbf{W}"'] & \mathbb{Y} \subset \mathbb{R}^r \arrow[d, shift left, "\mathbf{C}"] \\
\mathbb{T} \subset \mathbb{R}^\ell \arrow[r, "\mathbf{B}"] \arrow[u, shift right, "\mathbf{P}"'] & \mathbb{U} \subset \mathbb{R}^s \arrow[u, shift left, "\mathbf{Q}"]
\end{tikzcd}$$

Согласование проекций

Для получения согласованной модели в скрытом пространстве находится функция связи

$$\mathbf{U} = \mathbf{T}\mathbf{B}, \quad \mathbf{B} = \operatorname{diag}(\beta_k), \quad \beta_k = \boldsymbol{\nu}_k^{\mathsf{T}} \boldsymbol{\tau}_k / (\boldsymbol{\tau}_k^{\mathsf{T}} \boldsymbol{\tau}_k).$$

Финальная модель декодирования имеет вид

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^{\mathsf{T}} + \mathbf{E}_\mathbf{y} \approx \mathbf{T}\mathbf{B}\mathbf{Q}^{\mathsf{T}} + \mathbf{E}_\mathbf{y} = \mathbf{X}\mathbf{W}^*\mathbf{B}\mathbf{Q}^{\mathsf{T}} + \mathbf{E} = \mathbf{X}\boldsymbol{\Theta}^{\mathsf{T}} + \mathbf{E}_\mathbf{y},$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^*, \quad \text{где } \mathbf{W}^* = \mathbf{W}(\mathbf{P}^{\mathsf{T}}\mathbf{W})^{-1}.$$

Согласование зависимостей в задаче декодирования

Особенностью задачи является избыточность размерности пространств переменных $\mathbf{x}$ и $\mathbf{y}$. Требуется найти многообразия низкой размерности:

$$\begin{tikzcd}
\mathbb{X} \subset \mathbb{R}^n \arrow[r, "\mathbf{f}"] \arrow[d, shift right, "\varphi_{\mathbf{x}}"'] & \mathbb{Y} \subset \mathbb{R}^r \arrow[d, shift left, "\varphi_{\mathbf{y}}"] \\
\mathbb{T} \subset \mathbb{R}^\ell \arrow[r, "\mathbf{h}"] \arrow[u, shift right, "\psi_{\mathbf{x}}"'] & \mathbb{U} \subset \mathbb{R}^s \arrow[u, shift left, "\psi_{\mathbf{y}}"]
\end{tikzcd}$$

$\mathbb{T} \subset \mathbb{R}^\ell$ и $\mathbb{U} \subset \mathbb{R}^s$ *скрытые пространства* для $\mathbb{X} \in \mathbb{R}^n$ $(\ell \leqslant n)$ и $\mathbb{Y} \in \mathbb{R}^r (s \leqslant r)$, если существуют *функции кодирования* $\boldsymbol{\varphi}_{\mathbf{x}} : \mathbb{X} \to \mathbb{T}$, $\boldsymbol{\varphi}_{\mathbf{y}} : \mathbb{Y} \to \mathbb{U}$ и *декодирования* $\boldsymbol{\psi}_{\mathbf{x}} : \mathbb{T} \to \mathbb{X}$, $\boldsymbol{\psi}_{\mathbf{y}} : \mathbb{U} \to \mathbb{Y}$:

$$\text{для любого } \mathbf{x} \in \mathbb{X} \quad \text{существует } \mathbf{t} \in \mathbb{T} : \boldsymbol{\psi}_{\mathbf{x}}\bigl(\boldsymbol{\varphi}_{\mathbf{x}}(\mathbf{x})\bigr) = \boldsymbol{\psi}_{\mathbf{x}}(\mathbf{t}) = \mathbf{x};$$

$$\text{для любого } \mathbf{y} \in \mathbb{Y} \quad \text{существует } \mathbf{u} \in \mathbb{U} : \boldsymbol{\psi}_{\mathbf{y}}\bigl(\boldsymbol{\varphi}_{\mathbf{y}}(\mathbf{y})\bigr) = \boldsymbol{\psi}_{\mathbf{y}}(\mathbf{u}) = \mathbf{y}.$$

Скрытые пространства $\mathbb{T}$ и $\mathbb{U}$ называются **согласованными**, если существует *функция связи* $\mathbf{h} : \mathbb{T} \to \mathbb{U}$:

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = \boldsymbol{\psi}_{\mathbf{y}}\Bigl(\mathbf{h}\bigl(\boldsymbol{\varphi}_{\mathbf{x}}(\mathbf{x})\bigr)\Bigr).$$

Функция согласования проекций

$$g : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}, \quad g(\boldsymbol{\tau}, \boldsymbol{\nu}) \to \max_{\boldsymbol{\varphi}_{\mathbf{x}}, \boldsymbol{\varphi}_{\mathbf{y}}, \mathbf{h}}$$

Согласованная модель проекции в скрытое пространство

Утверждение (Исаченко, 2017)

Вычисленные вектора $\boldsymbol{\tau}_k$ *и* $\boldsymbol{\nu}_k$ *с помощью итеративной процедуры обновления:*

$$\boldsymbol{\tau}_k := \frac{\mathbf{X}_k \mathbf{w}_k}{\|\mathbf{w}_k\|}, \quad \mathbf{w}_k := \mathbf{X}_k^{\mathsf{T}} \boldsymbol{\nu}_{k-1} / (\boldsymbol{\nu}_{k-1}^{\mathsf{T}} \boldsymbol{\nu}_{k-1});$$

$$\boldsymbol{\nu}_k := \frac{\mathbf{Y}_k \mathbf{c}_k}{\|\mathbf{c}_k\|}, \quad \mathbf{c}_k := \mathbf{Y}_k^{\mathsf{T}} \boldsymbol{\tau}_k / (\boldsymbol{\tau}_k^{\mathsf{T}} \boldsymbol{\tau}_k).$$

обладают максимальной ковариацией $\mathrm{cov}(\boldsymbol{\tau}, \boldsymbol{\nu})$.

Теорема (Исаченко, 2017)

В случае линейных функций декодирования $\varphi_{\mathbf{x}}(\mathbf{T}) = \mathbf{T}\mathbf{P}^{\mathsf{T}}$, $\varphi_{\mathbf{y}}(\mathbf{U}) = \mathbf{U}\mathbf{Q}^{\mathsf{T}}$ *и функции согласования* $g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \mathrm{cov}(\boldsymbol{\tau}, \boldsymbol{\nu})$ *параметры*

$$\boldsymbol{\Theta} = \mathbf{W}(\mathbf{P}^{\mathsf{T}}\mathbf{W})^{-1}\mathbf{B}\mathbf{Q}^{\mathsf{T}}$$

являются оптимальными для модели $\mathbf{F}(\mathbf{X}, \boldsymbol{\Theta})$.

Согласованная модель проекции в скрытое пространство

Principal Component Analysis

$$\varphi_{\mathbf{x}}(\mathbf{X}) = \underset{m\times n}{\mathbf{X}} \cdot \underset{n\times l}{\mathbf{P}}^{\mathsf{T}}, \quad \psi_{\mathbf{x}}(\mathbf{T}) = \underset{m\times l}{\mathbf{T}} \cdot \underset{l\times n}{\mathbf{P}}, \quad \mathbf{P}\mathbf{P}^{\mathsf{T}} = \mathbf{I}.$$

$$\mathbf{p} = \arg\max_{\|\mathbf{p}\|_2=1} g(\boldsymbol{\tau}) = \arg\max_{\|\mathbf{p}\|_2=1} [\mathrm{var}(\mathbf{X}\mathbf{p})],$$

Partial Least Squares/Canonical Correlation Analysis

$$\varphi_{\mathbf{x}}(\mathbf{X}) = \mathbf{X}\mathbf{W}, \quad \varphi_{\mathbf{y}}(\mathbf{Y}) = \mathbf{Y}\mathbf{C},$$

$$\psi_{\mathbf{x}}(\mathbf{T}) = \mathbf{T}\mathbf{P}^{\mathsf{T}}, \quad \psi_{\mathbf{y}}(\mathbf{U}) = \mathbf{U}\mathbf{Q}^{\mathsf{T}}.$$

$$g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \mathrm{cov}(\boldsymbol{\tau}, \boldsymbol{\nu}), \quad g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \mathrm{corr}(\boldsymbol{\tau}, \boldsymbol{\nu}).$$

Deep CCA

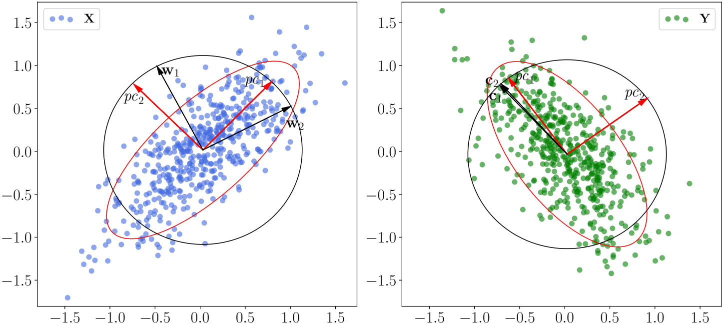
$$\max_{\|\mathbf{p}\|_2=\|\mathbf{q}\|_2=1} \left[\mathrm{corr}(\varphi_{\mathbf{x}}(\mathbf{X}) \cdot \mathbf{p}, \varphi_{\mathbf{y}}(\mathbf{Y}) \cdot \mathbf{q})^2\right] =$$

$$= \max_{\mathbf{p},\mathbf{q}} \frac{\mathbf{p}^{\mathsf{T}} \varphi_{\mathbf{x}}(\mathbf{X})^{\mathsf{T}} \varphi_{\mathbf{y}}(\mathbf{Y})\mathbf{q}}{\sqrt{\mathbf{p}^{\mathsf{T}} \varphi_{\mathbf{x}}(\mathbf{X})^{\mathsf{T}} \varphi_{\mathbf{x}}(\mathbf{X}, \mathbf{W}_{\mathbf{x}})\mathbf{p}}\sqrt{\mathbf{q}^{\mathsf{T}} \varphi_{\mathbf{y}}(\mathbf{Y})^{\mathsf{T}} \varphi_{\mathbf{y}}(\mathbf{Y})\mathbf{q}}}.$$

Пример согласованной проекции в скрытое пространство

Исходные переменные $\mathbf{x}_i \sim \mathcal{N}(0, \mathbf{\Sigma})$.

Целевые переменные $\mathbf{y}_i$ линейно зависят от pc_2 и не зависят от pc_1.



	LR	PCA	PLS	CCA
MSE	0.01	0.24	0.13	0.13

Согласование проекций матриц **X** и **Y** находит оптимальное скрытое представление, отклоняя вектора $\mathbf{w}_k$ и $\mathbf{c}_k$ от направления главных компонент.

Суперпозиция моделей декодирования сигналов

Пусть $\mathbf{f}_1(\mathbf{x}_1, \mathbf{\Theta}_1)$, $\mathbf{f}_2(\mathbf{x}_2, \mathbf{\Theta}_2)$ – линейные модели декодирования сигналов.

Утверждение (Исаченко, 2021)

Пусть модель декодирования является аддитивной суперпозицией линейных моделей:

$$\mathbf{y} = \mathbf{f}_1(\mathbf{x}_1, \mathbf{\Theta}_1) + \mathbf{f}_2(\mathbf{x}_2, \mathbf{\Theta}_2) + \varepsilon_{\mathbf{y}} = \mathbf{\Theta}_1\mathbf{x}_1 + \mathbf{\Theta}_2\mathbf{x}_2 + \varepsilon_{\mathbf{y}}.$$

Тогда оптимальные параметры

$$\mathbf{\Theta}_1 = (\mathbf{X}_1^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{X}_1)^{-1}\mathbf{X}_1^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{Y},$$
$$\mathbf{\Theta}_2 = (\mathbf{X}_2^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_1}\mathbf{X}_2)^{-1}\mathbf{X}_2^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_1}\mathbf{Y},$$

где $\mathbf{M}_{\mathbf{X}_1} = \mathbf{I} - \mathbf{X}_1(\mathbf{X}_1^{\mathsf{T}}\mathbf{X}_1)^{-1}\mathbf{X}_1^{\mathsf{T}}$, $\mathbf{M}_{\mathbf{X}_2} = \mathbf{I} - \mathbf{X}_2(\mathbf{X}_2^{\mathsf{T}}\mathbf{X}_2)^{-1}\mathbf{X}_2^{\mathsf{T}}$.

Теорема (Исаченко, 2021)

Ошибка аддитивной суперпозиции линейных моделей декодирования не превышает ошибки отдельной модели:

$$\mathcal{L}_{\text{sup}}(\mathbf{\Theta}_1^*, \mathbf{\Theta}_2^*, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leqslant \mathcal{L}(\mathbf{\Theta}_1, \mathbf{X}_1, \mathbf{Y}),$$
$$\mathcal{L}_{\text{sup}}(\mathbf{\Theta}_1^*, \mathbf{\Theta}_2^*, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leqslant \mathcal{L}(\mathbf{\Theta}_2, \mathbf{X}_2, \mathbf{Y}).$$

Суперпозиция моделей декодирования сигналов

Пусть $\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\Theta}_1)$, $\mathbf{f}_2(\mathbf{x}_2, \boldsymbol{\Theta}_2)$ – линейные модели декодирования сигналов.

Утверждение (Исаченко, 2021)

Пусть модель декодирования является аддитивной суперпозицией линейных моделей:

$$\mathbf{y} = \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\Theta}_1) + \mathbf{f}_2(\mathbf{x}_2, \boldsymbol{\Theta}_2) + \varepsilon_{\mathbf{y}} = \boldsymbol{\Theta}_1 \mathbf{x}_1 + \boldsymbol{\Theta}_2 \mathbf{x}_2 + \varepsilon_{\mathbf{y}}.$$

Оптимальная подматрица $\boldsymbol{\Theta}_2$ является решением задачи регрессии

$$\|\mathbf{Y}_1 - \mathbf{X}_{21}\boldsymbol{\Theta}_2\|^2 \to \min_{\boldsymbol{\Theta}_2},$$

где $\mathbf{Y}_1 = \mathbf{M}_{\mathbf{X}_1}\mathbf{Y}$, $\mathbf{X}_{21} = \mathbf{M}_{\mathbf{X}_1}\mathbf{X}_2$.

Теорема (Исаченко, 2021)

Пусть выполнены следующие условия

$$\mathbf{Y} \neq \mathbf{P}_{\mathbf{X}_2}\mathbf{Y}, \quad \mathbf{X}_1 \neq \mathbf{P}_{\mathbf{X}_2}\mathbf{X}_1, \quad \mathbf{Y}^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{X}_1 \neq \mathbf{0}.$$

Тогда выполнено строгое неравенство

$$\mathcal{L}_{\text{sup}}(\boldsymbol{\Theta}_1^*, \boldsymbol{\Theta}_2^*, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) < \mathcal{L}(\boldsymbol{\Theta}_2, \mathbf{X}_2, \mathbf{Y}).$$

Нелинейные методы согласования скрытого пространства

Функции кодирования и декодирования являются глубокими нейросетями:

$$
\begin{aligned}
\mathbf{T} &= \varphi_{\mathbf{x}}(\mathbf{X}) = \mathbf{W}_x^L \sigma(\dots \sigma(\mathbf{W}_x^2 \sigma(\mathbf{X}\mathbf{W}_x^1)) \dots) \\
\mathbf{U} &= \varphi_{\mathbf{y}}(\mathbf{Y}) = \mathbf{W}_y^L \sigma(\dots \sigma(\mathbf{W}_y^2 \sigma(\mathbf{Y}\mathbf{W}_y^1)) \dots) \\
\mathbf{X} &= \psi_{\mathbf{x}}(\mathbf{T}) = \mathbf{W}_t^L \sigma(\dots \sigma(\mathbf{W}_t^2 \sigma(\mathbf{T}\mathbf{W}_t^1)) \dots) \\
\mathbf{Y} &= \psi_{\mathbf{y}}(\mathbf{U}) = \mathbf{W}_u^L \sigma(\dots \sigma(\mathbf{W}_u^2 \sigma(\mathbf{U}\mathbf{W}_u^1)) \dots)
\end{aligned}
$$

Согласование проекций

$$
g(\mathbf{T}, \mathbf{U}) \to \max_{\mathbf{W}}, \quad \mathbf{W} = \{\mathbf{W}_x^i, \mathbf{W}_y^i, \mathbf{W}_t^i, \mathbf{W}_u^i\}_{i=1}^L.
$$

Градиент функции согласования $g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \mathrm{corr}(\boldsymbol{\tau}, \boldsymbol{\nu})$ имеет вид

$$
\frac{\partial g(\boldsymbol{\tau}, \boldsymbol{\nu})}{\partial \boldsymbol{\tau}} = \frac{1}{\ell - 1}\left(\boldsymbol{\Sigma}_1^{-1/2}\mathbf{U}\mathbf{V}^{\mathsf{T}}\boldsymbol{\Sigma}_2^{-1/2}\mathbf{U} - \boldsymbol{\Sigma}_1^{-1/2}\mathbf{U}\mathbf{D}\mathbf{V}^{\mathsf{T}}\boldsymbol{\Sigma}_1^{-1/2}\right),
$$

где $\mathbf{U}, \mathbf{D}, \mathbf{V} = \mathrm{SVD}(\boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_2^{-1/2}$, $\boldsymbol{\Sigma}_1 = \frac{1}{\ell-1}\mathbf{T}\mathbf{T}^{\mathsf{T}}$, $\boldsymbol{\Sigma}_2 = \frac{1}{\ell-1}\mathbf{U}\mathbf{U}^{\mathsf{T}}$, $\boldsymbol{\Sigma}_{12} = \frac{1}{\ell-1}\mathbf{T}\mathbf{U}^{\mathsf{T}}$.

Выбор признаков в задаче декодирования

$\mathbf{X} = [\chi_1, \ldots, \chi_n] \in \mathbb{R}^{m \times n}$ – матрица исходных переменных;
$\mathbf{Y} = [\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_r] \in \mathbb{R}^{m \times r}$ – матрица целевых переменных.
Требуется найти бинарный вектор $\mathbf{a} = \{0, 1\}^n$, компоненты – индикаторы выбранных признаков.

Функция ошибки отбора признаков

$$\mathbf{z} = \arg\min_{\mathbf{z}' \in [0,1]^n} S(\mathbf{z}', \mathbf{X}, \mathbf{Y}), \quad a_j = [z_j > \tau].$$

Функция ошибки модели декодирования

$$\mathcal{L}(\boldsymbol{\Theta}_\mathbf{a}, \mathbf{X}_\mathbf{a}, \mathbf{Y}) = \left\| \mathbf{Y} - \mathbf{X}_\mathbf{a} \boldsymbol{\Theta}_\mathbf{a}^\mathsf{T} \right\|_2^2 \to \min_{\boldsymbol{\Theta}_\mathbf{a}}, \quad \mathbf{X}_\mathbf{a} = \{\chi_j : a_j = 1, \, j = 1, \ldots, n\}$$

Задача квадратичного программирования

$$S(\mathbf{z}, \mathbf{X}, \boldsymbol{\nu}) = (1 - \alpha) \cdot \underbrace{\mathbf{z}^\mathsf{T} \mathbf{Q} \mathbf{z}}_{\mathrm{Sim}(\mathbf{X})} - \alpha \cdot \underbrace{\mathbf{b}^\mathsf{T} \mathbf{z}}_{\mathrm{Rel}(\mathbf{X}, \boldsymbol{\nu})} \to \min_{\substack{\mathbf{z} \geqslant \mathbf{0}_n \\ \mathbf{1}_n^\mathsf{T} \mathbf{z} = 1}}.$$

$\mathbf{z} \in [0, 1]^n$ – значимость признаков;
$\mathbf{Q} = \left[\left| \mathrm{corr}(\chi_i, \chi_j) \right| \right]_{i,j=1}^n \in \mathbb{R}^{n \times n}$ – матрица парных взаимодействий признаков;
$\mathbf{b} = \left[\left| \mathrm{corr}(\chi_i, \boldsymbol{\nu}) \right| \right]_{i=1}^n \in \mathbb{R}^n$ – вектор релевантностей признаков к целевой переменной.

Выбор признаков с помощью квадратичного программирования

Задача квадратичного программирования

$$S(\mathbf{z}, \mathbf{X}, \boldsymbol{\nu}) = (1-\alpha)\cdot \underbrace{\mathbf{z}^{\mathsf{T}}\mathbf{Q}\mathbf{z}}_{\mathrm{Sim}(\mathbf{X})} - \alpha \cdot \underbrace{\mathbf{b}^{\mathsf{T}}\mathbf{z}}_{\mathrm{Rel}(\mathbf{X},\boldsymbol{\nu})} \to \min_{\substack{\mathbf{z}\geqslant \mathbf{0}_n \\ \mathbf{1}_n^{\mathsf{T}}\mathbf{z}=1}}.$$

Теорема (Исаченко, 2018)

Пусть матрица парных взаимодействий признаков $\hat{\mathbf{Q}}$ *получена полуопределенной релаксацией исходной матрицы* $\mathbf{Q}$*:*

$$\hat{\mathbf{Q}} = \mathbf{Q} - \lambda_{\min}(\mathbf{Q})\mathbf{I}.$$

Тогда задача выбора признаков с помощью квадратичного программирования имеет единственный глобальный минимум.

Агрегирование релевантностей по целевым векторам (RelAgg)

$$\mathbf{b} = \left[|\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\nu})|\right]_{i=1}^{n} \to \mathbf{b} = \left[\sum_{k=1}^{r} |\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\nu}_k)|\right]_{i=1}^{n}.$$

Недостаток: нет учёта зависимостей в целевом пространстве матрицы $\mathbf{Y}$.

Выбор признаков в задаче декодирования

Симметричный учёт значимостей (SymImp)

Штрафуем коррелированные целевые вектора с помощью $\mathrm{Sim}(\mathbf{Y})$:

$$S(\mathbf{z}, \mathbf{X}, \mathbf{Y}) = \alpha_1 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{Q}_x \mathbf{z}_x}_{\mathrm{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{B} \mathbf{z}_y}_{\mathrm{Rel}(\mathbf{X},\mathbf{Y})} + \alpha_3 \cdot \underbrace{\mathbf{z}_y^{\mathsf{T}} \mathbf{Q}_y \mathbf{z}_y}_{\mathrm{Sim}(\mathbf{Y})} \to \min_{\substack{\mathbf{z}_x \geqslant \mathbf{0}_n,\, \mathbf{1}_n^{\mathsf{T}} \mathbf{z}_x = 1 \\ \mathbf{z}_y \geqslant \mathbf{0}_r,\, \mathbf{1}_r^{\mathsf{T}} \mathbf{z}_y = 1}},$$

$$\mathbf{Q}_x = \left[\left|\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j)\right|\right]_{i,j=1}^{n}, \ \mathbf{Q}_y = \left[\left|\mathrm{corr}(\boldsymbol{\nu}_i, \boldsymbol{\nu}_j)\right|\right]_{i,j=1}^{r}, \ \mathbf{B} = \left[\left|\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\nu}_j)\right|\right]_{\substack{i=1,\dots,n, \\ j=1,\dots,r}}$$

$$\alpha_1 + \alpha_2 + \alpha_3 = 1, \quad \alpha_i \geqslant 0.$$

SymImp штрафует коррелированные целевые вектора, которые в меньшей мере объясняются признаками.

Асимметричный учёт значимостей (AsymImp)

$$\alpha_1 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{Q}_x \mathbf{z}_x}_{\mathrm{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\left(\mathbf{z}_x^{\mathsf{T}} \mathbf{B} \mathbf{z}_y - \mathbf{b}^{\mathsf{T}} \mathbf{z}_y\right)}_{\mathrm{Rel}(\mathbf{X},\mathbf{Y})} + \alpha_3 \cdot \underbrace{\mathbf{z}_y^{\mathsf{T}} \mathbf{Q}_y \mathbf{z}_y}_{\mathrm{Sim}(\mathbf{Y})} \to \min_{\substack{\mathbf{z}_x \geqslant \mathbf{0}_n,\, \mathbf{1}_n^{\mathsf{T}} \mathbf{z}_x = 1 \\ \mathbf{z}_y \geqslant \mathbf{0}_r,\, \mathbf{1}_r^{\mathsf{T}} \mathbf{z}_y = 1}}.$$

При $b_j = \max\limits_{i=1,\dots n} [\mathbf{B}]_{i,j}$ коэффициенты при $\mathbf{z}_y$ в $\mathrm{Rel}(\mathbf{X}, \mathbf{Y})$ неотрицательны.

Выбор признаков в задаче декодирования

$$\alpha_1 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{Q}_x \mathbf{z}_x}_{\mathrm{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{B} \mathbf{z}_y}_{\mathrm{Rel}(\mathbf{X},\mathbf{Y})} \to \min_{\substack{\mathbf{z}_x \geqslant \mathbf{0}_n \\ \mathbf{1}_n^{\mathsf{T}} \mathbf{z}_x = 1}}; \quad \alpha_3 \cdot \underbrace{\mathbf{z}_y^{\mathsf{T}} \mathbf{Q}_y \mathbf{z}_y}_{\mathrm{Sim}(\mathbf{Y})} + \alpha_2 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{B} \mathbf{z}_y}_{\mathrm{Rel}(\mathbf{X},\mathbf{Y})} \to \min_{\substack{\mathbf{z}_y \geqslant \mathbf{0}_r \\ \mathbf{1}_r^{\mathsf{T}} \mathbf{z}_y = 1}}.$$

Минимаксный подход (MinMax / MaxMin)

$$S(\mathbf{z}, \mathbf{X}, \mathbf{Y}) = \min_{\substack{\mathbf{z}_x \geqslant \mathbf{0}_n \\ \mathbf{1}_n^{\mathsf{T}} \mathbf{z}_x = 1}} \max_{\substack{\mathbf{z}_y \geqslant \mathbf{0}_r \\ \mathbf{1}_r^{\mathsf{T}} \mathbf{z}_y = 1}} \left(\text{or} \max_{\substack{\mathbf{z}_y \geqslant \mathbf{0}_r \\ \mathbf{1}_r^{\mathsf{T}} \mathbf{z}_y = 1}} \min_{\substack{\mathbf{z}_x \geqslant \mathbf{0}_n \\ \mathbf{1}_n^{\mathsf{T}} \mathbf{z}_x = 1}} \right) \left[\alpha_1 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{Q}_x \mathbf{z}_x}_{\mathrm{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^{\mathsf{T}} \mathbf{B} \mathbf{z}_y}_{\mathrm{Rel}(\mathbf{X},\mathbf{Y})} - \alpha_3 \cdot \underbrace{\mathbf{z}_y^{\mathsf{T}} \mathbf{Q}_y \mathbf{z}_y}_{\mathrm{Sim}(\mathbf{Y})} \right].$$

Теорема (Исаченко, 2018)

Для положительно определенных матриц $\mathbf{Q}_x$ и $\mathbf{Q}_y$ minmax и maxmin задачи достигают одинакового значения функционала $S(\mathbf{z}, \mathbf{X}, \mathbf{Y})$

Теорема (Исаченко, 2018)

Минимаксная задача эквивалентна задаче квадратичного программирования с $n + r + 1$ переменными.

Для получения выпуклой задачи применяется полуопределенная релаксация сдвига спектра.

Обобщение предложенных методов выбора признаков

Теорема (Исаченко, 2018)

В одномерном случае $r = 1$ предлагаемые методы выбора признаков SymImp, MinMax, MaxMin, AsymImp совпадают с исходной задачей минимизации функции ошибок $S(\mathbf{z}, \mathbf{X}, \mathbf{Y})$.

Метод	Критерий	Функция ошибки $S(\mathbf{z}, \mathbf{X}, \mathbf{Y})$
RelAgg	$\min\bigl[\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})\bigr]$	$\min\limits_{\mathbf{z}_x}\bigl[(1-\alpha)\cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{Q}_x\mathbf{z}_x - \alpha \cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{B}\mathbf{1}_r\bigr]$
SymImp	$\min\bigl[\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})\bigr]$	$\min\limits_{\mathbf{z}_x, \mathbf{z}_y}\bigl[\alpha_1 \cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{Q}_x\mathbf{z}_x - \alpha_2 \cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{B}\mathbf{z}_y + \alpha_3 \cdot \mathbf{z}_y^{\mathsf{T}}\mathbf{Q}_y\mathbf{z}_y\bigr]$
MinMax	$\min\bigl[\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})\bigr]$ $\max\bigl[\text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})\bigr]$	$\min\limits_{\mathbf{z}_x}\max\limits_{\mathbf{z}_y}\bigl[\alpha_1 \cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{Q}_x\mathbf{z}_x - \alpha_2 \cdot \mathbf{z}_x^{\mathsf{T}}\mathbf{B}\mathbf{z}_y - \alpha_3 \cdot \mathbf{z}_y^{\mathsf{T}}\mathbf{Q}_y\mathbf{z}_y\bigr]$
AsymImp	$\min\bigl[\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})\bigr]$ $\max\bigl[\text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})\bigr]$	$\min\limits_{\mathbf{z}_x, \mathbf{z}_y}\bigl[\alpha_1\mathbf{z}_x^{\mathsf{T}}\mathbf{Q}_x\mathbf{z}_x - \alpha_2\left(\mathbf{z}_x^{\mathsf{T}}\mathbf{B}\mathbf{z}_y - \mathbf{b}^{\mathsf{T}}\mathbf{z}_y\right) + \alpha_3\mathbf{z}_y^{\mathsf{T}}\mathbf{Q}_y\mathbf{z}_y\bigr]$

Оптимизация нелинейных моделей с помощью квадратичного программирования

Метод Ньютона:

$$\boldsymbol{\theta}^k = \boldsymbol{\theta}^{k-1} + \Delta\boldsymbol{\theta}^{k-1} = \boldsymbol{\theta}^{k-1} - \mathbf{H}^{-1}\nabla S(\boldsymbol{\theta}).$$

Обновление весов с выбором активных параметров:

$$\boldsymbol{\theta}^k_{\mathcal{A}} = \boldsymbol{\theta}^{k-1}_{\mathcal{A}} + \Delta\boldsymbol{\theta}^{k-1}_{\mathcal{A}}, \quad \boldsymbol{\theta}_{\mathcal{A}} = \{\theta_j : j \in \mathcal{A}\},$$
$$\boldsymbol{\theta}^k_{\bar{\mathcal{A}}} = \boldsymbol{\theta}^{k-1}_{\bar{\mathcal{A}}}, \quad \boldsymbol{\theta}_{\bar{\mathcal{A}}} = \{\theta_j : j \notin \mathcal{A}\}.$$

Функция ошибки выбора активных параметров:

$$\mathbf{a} = \arg\max_{\mathbf{a}' \in \{1,0\}^p} S(\mathbf{a}', \mathbf{X}, \mathbf{y}, \boldsymbol{\theta}) \Leftrightarrow \arg\min_{\mathbf{a}\in\mathbb{R}^p_+,\, \|\mathbf{a}\|_1=1} \left[\mathbf{a}^{\mathsf{T}}\mathbf{Q}\mathbf{a} - \alpha \cdot \mathbf{b}^{\mathsf{T}}\mathbf{a}\right].$$

Параметр θ_j для модели $f(\mathbf{x}, \boldsymbol{\theta})$ является **активным**, если $\mathbf{J}^{\mathsf{T}}(\mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) - \mathbf{y}) \neq 0$.

Функция ошибок для нелинейной регрессии:

$$\mathcal{L}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{y}) = \frac{1}{2}\|\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\theta})\|^2 = \frac{1}{2}\sum_{i=1}^{m}\big(y_i - f(\mathbf{x}_i, \boldsymbol{\theta})\big)^2,$$

Функция ошибок для логистической регрессии:

$$\mathcal{L}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{y}) = \sum_{i=1}^{m}\Big[y_i \log f(\mathbf{x}_i, \boldsymbol{\theta}) + (1 - y_i)\log\big(1 - f(\mathbf{x}_i, \boldsymbol{\theta})\big)\Big].$$

Оптимизация нелинейных моделей с помощью квадратичного программирования

Теорема (Исаченко, 2018)

Пусть модель $f(\mathbf{x}, \boldsymbol{\theta})$ *близка к линейной в окрестности точки* $\boldsymbol{\theta} + \Delta\boldsymbol{\theta}$

$$\mathbf{f}(\mathbf{X}, \boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \approx \mathbf{f}(\mathbf{X}, \boldsymbol{\theta}) + \mathbf{J} \cdot \Delta\boldsymbol{\theta},$$

где $\mathbf{J} \in \mathbb{R}^{m \times p}$ *является матрицей Якоби. Тогда вектор обновления* $\Delta\boldsymbol{\theta}$ *для функции ошибки нелинейной регрессии является решением задачи*

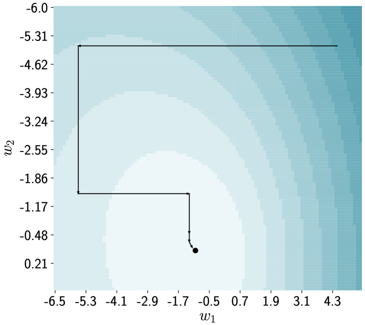
$$\|\mathbf{e} - \mathbf{F}\Delta\boldsymbol{\theta}\|_2^2 \to \min_{\Delta\boldsymbol{\theta} \in \mathbb{R}^p}, \quad \mathbf{e} = \mathbf{f} - \mathbf{y},\ \mathbf{F} = \mathbf{J}.$$

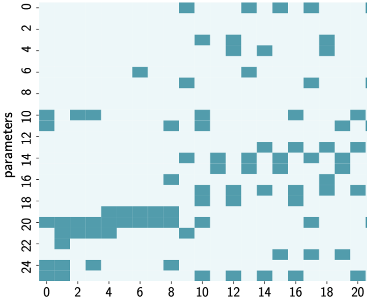
Теорема (Исаченко, 2018)

Пусть модель имеет вид $f(\mathbf{x}, \boldsymbol{\theta}) = \sigma(\mathbf{x}^{\mathsf{T}}\boldsymbol{\theta})$. *Вектор обновлений* $\Delta\boldsymbol{\theta}$ *для функции ошибки логистической регрессии является решением задачи*

$$\|\mathbf{e} - \mathbf{F}\Delta\boldsymbol{\theta}\|_2^2 \to \min_{\Delta\boldsymbol{\theta} \in \mathbb{R}^p}, \quad \mathbf{e} = \mathbf{R}^{-1/2}(\mathbf{y} - \mathbf{f}),\ \mathbf{F} = \mathbf{R}^{1/2}\mathbf{X}.$$

Оптимизация нелинейных моделей с помощью квадратичного программирования





Выборка	GD	Нестеров	ADAM	Ньютон	QPFS+Ньютон
Boston House Prices	27.2 ± 4.6 32.4 ± 5.6	46.0 ± 11.0 53.3 ± 11.5	35.4 ± 2.5 37.8 ± 7.0	22.1 ± 15.2 28.9 ± 13.6	20.9 ± 10.4 $\mathbf{24.5 \pm 9.4}$
Communities and Crime	48.0 ± 6.4 $47,5 \pm 6.5$	31.4 ± 2.8 32.9 ± 4.3	23.3 ± 3.7 $28,1 \pm 4.5$	18.3 ± 3.4 28.8 ± 3.6	26.7 ± 3.1 $\mathbf{28.4 \pm 3.0}$
Forest Fires	18.9 ± 0.4 $\mathbf{20.0 \pm 2.1}$	1.83 ± 0.4 20.2 ± 2.2	1.81 ± 0.6 $\mathbf{20.0 \pm 2.0}$	17.7 ± 0.4 20.6 ± 1.4	17.9 ± 0.4 20.2 ± 2.2
Residential Building	51.6 ± 17.7 53.7 ± 13.9	32.6 ± 19.5 34.1 ± 13.6	30.0 ± 24.8 34.1 ± 19.4	35.5 ± 24.7 35.0 ± 15.6	30.3 ± 10.7 $\mathbf{30.9 \pm 5.3}$

Внешние критерии качества решения задачи декодирования

Нормированное RMSE

Качество прогнозирования:

$$\mathrm{sRMSE}(\mathbf{Y}, \widehat{\mathbf{Y}}_{\mathbf{a}}) = \sqrt{\frac{\mathrm{MSE}(\mathbf{Y}, \widehat{\mathbf{Y}}_{\mathbf{a}})}{\mathrm{MSE}(\mathbf{Y}, \overline{\mathbf{Y}})}} = \frac{\|\mathbf{Y} - \widehat{\mathbf{Y}}_{\mathbf{a}}\|_2}{\|\mathbf{Y} - \overline{\mathbf{Y}}\|_2}, \quad \text{где} \quad \widehat{\mathbf{Y}}_{\mathbf{a}} = \mathbf{X}_{\mathbf{a}} \boldsymbol{\Theta}_{\mathbf{a}}^{\mathsf{T}}.$$

$\overline{\mathbf{Y}}$ — константный прогноз.

Мультикорреляция

Среднее значение коэффициента множественной корреляции:

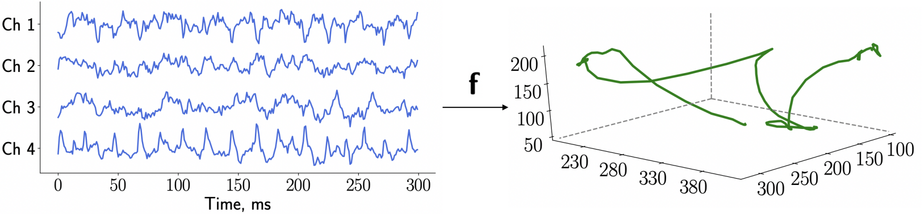
$$R^2 = \frac{1}{r} \mathrm{tr}\left(\mathbf{C}^{\mathsf{T}} \mathbf{R}^{-1} \mathbf{C}\right), \quad \mathbf{C} = [\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\nu}_j)]_{\substack{i=1,\dots,n \\ j=1,\dots,r}}, \; \mathbf{R} = [\mathrm{corr}(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j)]_{i,j=1}^{n}.$$

Байесовский информационный критерий

Компромисс между качеством предсказания и числом выбранных признаков $\|\mathbf{a}\|_0$:

$$\mathrm{BIC} = m \ln\left(\mathrm{MSE}(\mathbf{Y}, \widehat{\mathbf{Y}}_{\mathbf{a}})\right) + \|\mathbf{a}\|_0 \cdot \log m.$$

Задача декодирования сигналов электрокортикограммы

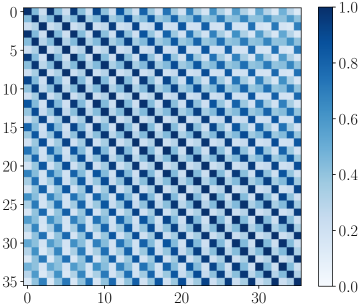


Заданы:
$\mathbf{X} \in \mathbb{R}^{m \times (32 \cdot 27)}$ – сигналы ECoG,
$\mathbf{Y} \in \mathbb{R}^{m \times 3k}$ – траектория движения руки, где

$$\mathbf{Y} = \begin{pmatrix} x_1 \; y_1 \; z_1 & \dots & x_k & y_k & z_k \\ x_2 \; y_2 \; z_2 & \dots & x_{k+1} & y_{k+1} & z_{k+1} \\ \dots & \dots & & \dots & \\ x_m \; y_m \; z_m & \dots & x_{m+k} & y_{m+k} & z_{m+k} \end{pmatrix}.$$

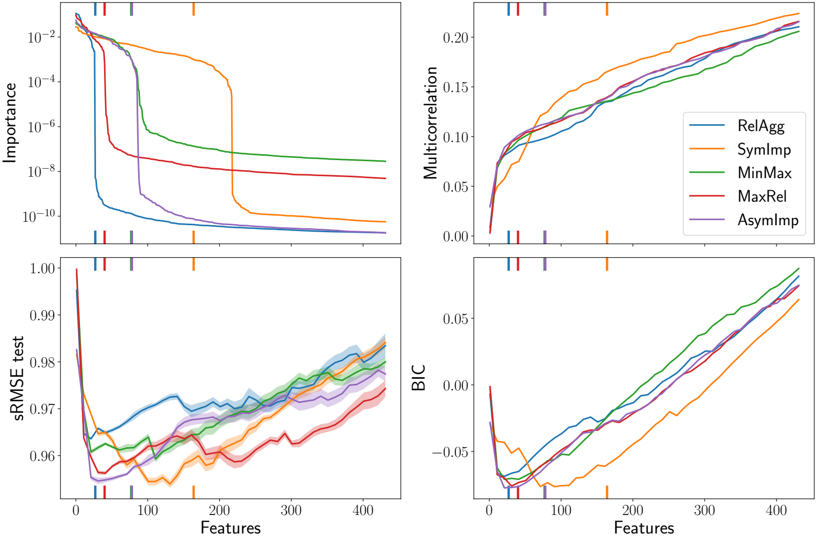
Столбцы матрицы $\mathbf{Y}$ сильно скоррелированы по временной оси.

`http://neurotycho.org`



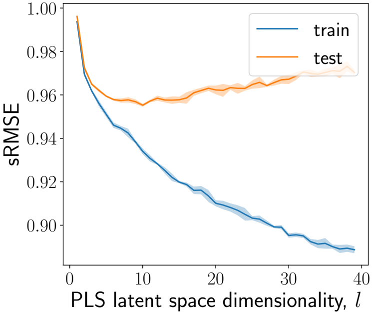
Матрица корреляций $\mathbf{Y}$

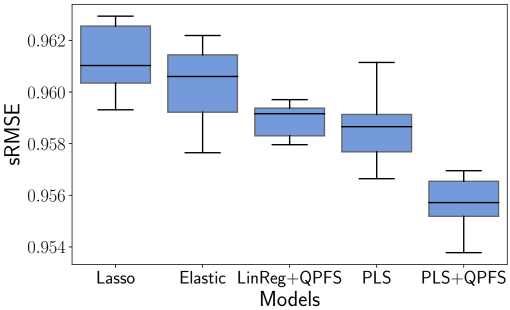
Анализ предложенных методов выбора признаков



Предложены методы выбора модели, имеющей меньшую ошибкой по отношению к базовому алгоритму.

Сравнение метода проекции в скрытое пространство с методами выбора признаков





	sRMSE	$\|\mathbf{a}\|_0$	Spearman ρ	ℓ_2
RelAgg	0.965 ± 0.002	26.8 ± 3.8	0.915 ± 0.016	0.145 ± 0.018
SymImp	0.961 ± 0.001	224.4 ± 9.0	0.910 ± 0.017	0.025 ± 0.002
MinMax	0.961 ± 0.002	101.0 ± 2.1	0.932 ± 0.009	0.059 ± 0.004
AsymImp	0.955 ± 0.001	85.8 ± 10.2	0.926 ± 0.011	0.078 ± 0.007

Результаты, выносимые на защиту

1. Исследована проблема снижения размерности сигналов в коррелированных пространствах высокой размерности. Предложены методы декодирования сигналов, учитывающие зависимости как в исходном, так и в целевом пространстве сигналов.

2. Доказаны теоремы об оптимальности предлагаемых методов декодирования сигналов. Предлагаемые методы выбирают согласованные модели в случае избыточной размерности описания данных.

3. Предложены методы выбора признаков, учитывающие зависимости как в исходном, так и в целевом пространстве. Предложенные методы доставляют устойчивые и адекватные решения в пространствах высокой размерности.

4. Предложены нелинейные методы согласования скрытых пространств для данных со сложноорганизованной целевой переменной.

5. Предложен ряд моделей для прогнозирования гетерогенных наборов сигналов для задачи построения нейрокомпьютерных интерфейсов.

Список работ автора по теме диссертации

Публикации ВАК

1. Isachenko R., Strijov V. Quadratic Programming Feature Selection for Multicorrelated Signal Decoding with Partial Least Squares *Expert Systems with Applications*, 2021, на рецензировании.
2. Исаченко Р.В., Яушев Ф.Р., Стрижов В.В. Модели согласования скрытого пространства в задаче прогнозирования // Системы и средства информатики, 31(1), 2021.
3. Isachenko R., Vladimirova M., Strijov V. Dimensionality Reduction for Time Series Decoding and Forecasting Problems. *DEStech Transactions on Computer Science and Engineering*, optim, 2018.
4. Isachenko R., Strijov V. Quadratic programming optimization for Newton method. *Lobachevskii Journal of Mathematics*, 39(9), 2018.
5. Isachenko R. et al. Feature Generation for Physical Activity Classification. *Artificial Intellegence and Decision Making*, 3, 2018.
6. Исаченко Р.В., Стрижов В. В. Метрическое обучение в задачах мультиклассовой классификации временных рядов *Информатика и её применения*, 10(2), 2016.

Выступления с докладом

1. Intelligent Data Processing Conference, 2020, Снижение размерности в задаче декодирования временных рядов.
2. Intelligent Data Processing Conference, 2018, Dimensionality reduction for multicorrelated signal decoding with projections to latent space.
3. Математические методы распознавания образов, 2017. Локальные модели для классификации объектов сложной структуры.
4. Intelligent Data Processing Conference, 2016. Multimodel forecasting multiscale time series in internet of things.
5. Ломоносов, 2016. Метрическое обучение в задачах мультиклассовой классификации временных рядов.